

Feedback Reduction for MIMO Broadcast Channel with Heterogeneous Fading

Jin-Hao Li and Hsuan-Jung Su

Graduate Institute of Communication Engineering

Department of Electrical Engineering

National Taiwan University, Taipei, Taiwan

Email: jinghaw2003@gmail.com, hjsu@cc.ee.ntu.edu.tw

Abstract—This paper considers feedback load reduction for multiuser multiple input multiple output (MIMO) broadcast channel where the users' channel distributions are not homogeneous. A cluster-based feedback scheme is proposed such that the range of possible signal-to-noise ratio (SNR) of the users are divided into several clusters according to the order statistics of the users' SNRs. Each cluster has a corresponding threshold, and the users compare their measured instantaneous SNRs with the thresholds to determine whether and how many bits they should use to feed back their instantaneous SNRs. If a user's instantaneous SNR is lower than a certain threshold, the user does not feed back. Feedback load reduction is thus achieved. For a given number of clusters, the sum rate loss using the cluster-based feedback scheme is investigated. Then the minimum number of clusters given a maximum tolerable sum rate loss is derived. Through simulations, it is shown that, when the number of users is large, full multiuser diversity can be achieved by the proposed feedback scheme, which is more efficient than the conventional schemes.

I. INTRODUCTION

Multiuser diversity can significantly improve system throughput when the users suffer channel fluctuations. The performance gain of multiuser diversity grows with the number of users when the scheduler performs maximum throughput scheduling [1]. For the broadcast channel, the dirty paper coding (DPC) [2] achieves a sum rate which was shown in [3] to have a growth rate $M \log \log(K)$, where M is the number of transmit antennas at the base station (BS) and K is the number of users in the system. However, this scheme has high complexity in encoding/decoding and is difficult to be implemented. Therefore, a suboptimal and low-complexity zero-forcing (ZF) beamforming technique was proposed in [4] which also achieves the optimal growth rate of the sum rate. The results both in DPC and ZF schemes were based on the full channel state information (CSI) assumption at the transmitter and thus the users are required to feedback perfect CSI to the BS. Although the optimal sum rate can be achieved, the feedback load will increase linearly with the number of users.

Various approaches have been proposed to limit the amount of feedback load and investigate sum rate loss. In [5][6], the quantized channel direction information (CDI) was used to characterize the sum rate loss when the ZF technique is considered. It was shown that the number of feedback bits of each user needs to be increased linearly with the transmission power

to maintain a constant sum rate loss. Another low-feedback-rate and practical scheme, orthogonal random beamforming (ORB) was proposed in [7]. In ORB, each user only feeds back the CSI and the beam index of its favorite beam to the BS. Therefore, the total amount of feedback can be reduced from MK CSI values to K CSI values. Besides, the sum rate loss is negligible when the number of users is large [8]. In an effort to further reduce feedback load, a threshold based mechanism was proposed in [9] such that a user does not feed back when its CSI is below the threshold. In that work, the design of threshold does not take the scheduling algorithm into account. In [10], multiple thresholds were proposed. The design of multiple thresholds was based on the order statistics of the signal-to-interference-noise-ratio (SINR) assuming that greedy sum rate scheduling was performed in a *homogeneous* network where the users' channel gain distributions are the same. Exhaustive search was used to obtain the thresholds, which resulted in high computational complexity.

In this work, the closed form solution of the multiple thresholds in [10] is found, and a more realistic assumption on the channel distributions of the users is considered. We assume that the distributions of the users' signal-to-noise ratios (SNRs) are independent and non-identical. Since the computational complexity is too high to consider all the users' non-identical channel statistics, the statistics of the users' channels are divided into multiple clusters. The statistics in each cluster are used to calculate the corresponding threshold. Each cluster corresponds to an SNR range and is quantized with a few bits for differentiating the users' SNR levels falling in the same cluster. In addition, the sum rate loss due to setting thresholds is investigated. The performance of the proposed scheme is compared with that of the conventional feedback scheme and single threshold feedback scheme [9] in terms of sum rate, feedback load and efficiency. Through simulations, it is shown that the proposed cluster-based feedback scheme is more efficient than conventional feedback schemes and achieves higher sum rate than the single threshold feedback scheme.

II. SYSTEM MODEL

The multiuser multiple-input multiple-output (MIMO) down-link system is considered. The BS is equipped with M antennas and there are K users, each having N receive antennas.

According to the ORB strategy for multiuser transmission, the BS uses a precoding matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M]$ to simultaneously transmit signals, where $\mathbf{w}_i \in \mathbb{C}^{M \times 1}$, $i = 1, 2, \dots, M$, are random orthogonal vectors generated from isotropic distribution [11]. Let $\mathbf{s} = \sum_m \mathbf{w}_m s_m$ be the $M \times 1$ vector of the transmitted signal, where s_m is the m th transmitted symbol. The received signal for user k is given by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{W} \mathbf{s} + \mathbf{n}_k \quad (1)$$

where \mathbf{H}_k denotes the channel matrix between the BS and user k . The elements of the channel matrix \mathbf{H}_k are assumed to be independent identically distributed (i.i.d.) complex Gaussian random variables with zero mean and variance σ_k^2 . The noise term for user k , denoted \mathbf{n}_k , is modeled as i.i.d. zero mean complex Gaussian with covariance matrix $\mathbb{E}[\mathbf{n}_k \mathbf{n}_k^H] = \sigma_N^2 \mathbf{I}$, $\forall k$, where \mathbb{E} denotes the expectation operation and $(\cdot)^H$ represents the transpose conjugate.

User k uses the ZF receiver [12] to perform channel inversion to the received signal \mathbf{y}_k . Thus, the received signal after ZF receiver is given by

$$(\mathbf{H}_k \mathbf{W})^\dagger \mathbf{y}_k = \mathbf{s} + (\mathbf{H}_k \mathbf{W})^\dagger \mathbf{n}_k \quad (2)$$

where $\mathbf{H}^\dagger = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H$ is the pseudo-inverse of \mathbf{H} . Under the equal power assumption, let the transmit energy of each antenna be $\frac{P}{M}$, where P is the total transmit power at the BS. The SNR for the k th user at the m th spatial channel is denoted by $X_{m,k}$

$$X_{m,k} = \frac{P/M}{\sigma_N^2 [((\mathbf{H}_k \mathbf{W})^H (\mathbf{H}_k \mathbf{W}))^{-1}]_m} \quad m = 1, 2, \dots, M \quad (3)$$

where $[\mathbf{A}]_m$ denotes the m th diagonal element of matrix \mathbf{A} . Assuming $N \geq M$, it is well known that $X_{m,k}$ is a chi-square random variable with $2(N - M + 1)$ degrees of freedom [13][14]. Then the probability density function (PDF) of $X_{m,k}$, $\forall m$, can be expressed as

$$f_k(x) = \frac{\lambda_k^{N-M+1} x^{N-M} e^{-\lambda_k x}}{(N-M)!} \quad (4)$$

where $\lambda_k = \frac{M\sigma_N^2}{P\sigma_k^2}$. For simplicity, we drop the index m , denote $X_{m,k}$ as X_k , and restrict our analysis for the case $N = M$. Therefore, the distribution of X_k becomes the exponential distribution with parameter λ_k .

Let $X_{(1)}, X_{(2)}, \dots, X_{(K)}$, be the order statistics of independent continuous random variables X_1, X_2, \dots, X_K , with the PDF (4) in decreasing order, i.e., $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(K)}$. The cumulative distribution function (CDF) of the largest order statistics $X_{(1)}$ can be shown as

$$F_{X_{(1)}}(x) = \prod_{i=1}^K (1 - e^{-\lambda_i x}) \quad (5)$$

and its corresponding PDF can be expressed as

$$f_{X_{(1)}}(x) = \frac{1}{(K-1)!} \sum_{\mathbf{T}} F_{i_1}(x) \cdots F_{i_{K-1}}(x) f_{i_K}(x) \quad (6)$$

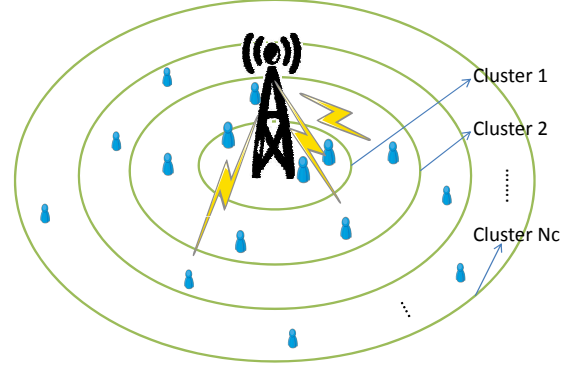


Fig. 1. Cluster-based Model.

where $\sum_{\mathbf{T}}$ denotes the summation over all $K!$ permutations (i_1, i_2, \dots, i_K) of $(1, 2, \dots, K)$. Applying the maximum sum rate scheduling algorithm, the sum rate of the system can be written as follows:

$$\begin{aligned} R &= \mathbb{E} \left\{ \sum_{m=1}^M \log_2 \left(1 + \max_{1 \leq k \leq K} X_{m,k} \right) \right\} \\ &= M \int_0^\infty \log_2(1+x) f_{X_{(1)}}(x) dx. \end{aligned} \quad (7)$$

In order to achieve the sum rate described in (7), each user should feed back the index of the precoding vector on which it sees the highest SNR, and the corresponding SNR value.

III. CLUSTER-BASED FEEDBACK MODEL

The transmission and feedback procedure can be described as follows. From the previous channel condition or location information feedbacks from the users, the BS derives the users' mean SNR. The BS then groups the users' mean SNRs into N_c clusters according to their magnitudes. The mean SNRs in each cluster are similar in quantity and are used to derive one SNR threshold, denoted $r_{c,i}$ for cluster i . Note that derivation of the users' mean SNRs and the cluster thresholds is done periodically. Derivation of the users' mean SNRs does not need to be very accurate. The BS broadcasts periodically the threshold set $\{r_{c,1} \geq r_{c,2} \geq \dots \geq r_{c,N_c}\}$ to the users.

At every feedback instant, each user compares its measured instantaneous SNR with the thresholds to determine which cluster its instantaneous SNR belongs to, and feeds back to the BS the cluster index and the quantization bits of the instantaneous SNR for this cluster. When a user's SNR is smaller than r_{c,N_c} , the user does not feedback to the BS. The proposed procedure takes a little downlink bandwidth for the BS to periodically broadcast the threshold set, in exchange of reduction of the uplink feedback bandwidth.

To derive the thresholds, the overall statistics of users are divided into multiple clusters according to the mean SNRs $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_K}$. The means of the random variables X_1, X_2, \dots, X_K are ranked with decreasing order as $\frac{1}{\lambda_{(1)}} \geq \frac{1}{\lambda_{(2)}} \geq \dots \geq \frac{1}{\lambda_{(K)}}$ and uniformly divided into N_c clusters with size $L = K/N_c$ as $\left\{ \frac{1}{\lambda_{(1)}}, \frac{1}{\lambda_{(2)}}, \dots, \frac{1}{\lambda_{(L)}} \right\}, \dots$,

$\left\{ \frac{1}{\lambda_{(L(N_c-1)+1)}}, \frac{1}{\lambda_{(L(N_c-1)+1)}}, \dots, \frac{1}{\lambda_{(K)}} \right\}$. To simplify the notation, we let random variable Y_m^n represent the m th statistic in the n th cluster. Then, Y_m^n is exponentially distributed with mean value $\frac{1}{\lambda_{(L(n-1)+m)}}$. The random variables in cluster n are $\{Y_1^n, Y_2^n, \dots, Y_L^n\}$, and let $Y_{(1)}^n \geq Y_{(2)}^n \geq \dots \geq Y_{(L)}^n$ be the order statistics of them. For a measured instantaneous SNR y_i at user i , we say that the rank of y_i in cluster n is d if $\underbrace{Y_{(1)}^n \geq Y_{(2)}^n \geq Y_{(3)}^n \dots \geq y_i}_{(d-1) \text{ variables}} \geq \underbrace{Y_{(d+1)}^n \dots \geq Y_{(K)}^n}_{(L-d) \text{ variables}}$. In the following, we will discuss how to derive the threshold in each cluster for heterogenous and homogeneous channel distributions to reduce the feedback load.

A. Heterogenous Case

1) *Cluster-based Type-I*: With maximum sum rate scheduling, a user will be scheduled when its SNR on a particular precoding vector is the highest among the users. On the other hand, if a user has a low SNR, it is unlikely to be scheduled. For this user, feeding back CSI is wasteful of the uplink radio resource. The threshold of each cluster is designed according to the probability of a user's measured instantaneous SNR being a particular rank in that cluster. Let $P_{m,n}^i(r)$ denote the probability that user m is ranked n among the L users in cluster i when its instantaneous SNR is r .

$$\begin{aligned} P_{m,n}^i(r) &= P\{Y_m^i = Y_{(n)}^i | Y_m^i = r\} \\ &= \sum_{\mathbf{S}} P\{\underbrace{Y_{t_1}^i \geq Y_{t_2}^i \geq \dots \geq Y_{t_{n-1}}^i}_{(n-1) \text{ variables}} \geq r \geq \underbrace{Y_{t_{n+1}}^i \dots \geq Y_{t_L}^i}_{(L-n) \text{ variables}}\} \\ &= \frac{\sum_{\mathbf{S}} e^{-\sum_{a=1}^{n-1} \lambda_{(L(i-1)+t_a)} r} \prod_{b=n+1}^L (1 - e^{-\lambda_{(L(i-1)+t_b)} r})}{(n-1)!(N_c - n)!} \end{aligned}$$

where $\sum_{\mathbf{S}}$ denotes the summation over all permutations $(t_1, \dots, t_{n-1}, t_{n+1}, \dots, t_L)$ of $(1, 2, \dots, m-1, m+1, \dots, L)$ in the cluster i . For example, when the instantaneous SNR of user t in cluster 1 is infinity, the rank of user t among the L users in cluster 1 is one with probability one, i.e., $P_{t,1}^1(\infty) = P\{Y_t^1 = Y_{(1)}^1 | Y_t^1 = \infty\} = 1$. Being valid conditional probabilities, the $P_{m,n}^i(r)$'s satisfy

$$\sum_{n=1}^L P_{m,n}^i(r) = 1, \quad m = 1, 2, \dots, L. \quad (8)$$

The most probable rank of user m in cluster i when its instantaneous SNR is r can be obtained by

$$\hat{n} = \arg \max_{n \in \{1, 2, \dots, L\}} P_{m,n}^i(r). \quad (9)$$

Let $Q_{m,n}^i$ be defined such that $P_{m,n}^i(Q_{m,n}^i) = P_{m,n+1}^i(Q_{m,n}^i)$. It can be seen that when the instantaneous SNR of user m falls in the range $[Q_{m,n}^i, Q_{m,n-1}^i)$, the most probable rank of user m in cluster i is n . For maximum sum rate scheduling, we let only the users who are most likely to

be rank one in each cluster to feedback. Therefore, a threshold for each cluster i is set as

$$r_{c,i} = \max\{Q_{1,1}^i, Q_{2,1}^i, \dots, Q_{L,1}^i\} \quad i = 1, 2, \dots, N_c \quad (10)$$

which satisfies the following inequality

$$r_{c,1} \geq r_{c,2} \geq \dots \geq r_{c,N_c}. \quad (11)$$

To reduce the computational complexity, another simple and low complexity method is proposed in the next section.

2) *Cluster-based Type-II*: We approximate by assuming that the random variables $Y_m^i, m = 1, 2, \dots, L$ in cluster i have the same exponential distribution with mean $\mu_{c,i}$ obtained by

$$\mu_{c,i} = \frac{\sum_{m=1}^L \lambda_{(L(i-1)+m)}}{L}. \quad (12)$$

Then, $P_{m,n}^i(r)$ is the same for all users in a cluster, and can be obtained by

$$P_{m,n}^i(r) = \frac{(L-1)! \exp(-\mu_{c,i} r)^{n-1} (1 - \exp(-\mu_{c,i} r))^{L-n}}{(n-1)!(L-n)!}. \quad (13)$$

With this approximation, the closed-form solution of $r_{c,i}$ can be derived as

$$r_{c,i} = Q_{1,1}^i = \frac{1}{\mu_{c,i}} \ln L, \quad i = 1, 2, \dots, N_c. \quad (14)$$

B. Homogeneous Case

The homogeneous case can be viewed as a special case of the cluster-based type-II scenario using only one cluster ($N_c = 1, K = L$). The mean values of the random variables $Y_m^1, m = 1, 2, \dots, K$ are the same, i.e., $1/\lambda = 1/\lambda_1 = \dots = 1/\lambda_K$. As in [10], multiple thresholds are set according to the most probable rank as

$$r_{c,p} = Q_{1,p}^1 = \frac{1}{\lambda} \ln \left(\frac{K}{p} \right), \quad p = 1, 2, \dots, N_c. \quad (15)$$

For all cases, at each feedback instant, when a user's instantaneous SNR is greater than the smallest threshold r_{c,N_c} , the user needs to feed back to the BS using $B_C = \log_2 \lceil N_c \rceil$ bits to indicate which region (between two adjacent thresholds) its instantaneous SNR belongs to. In addition, in order to differentiate the users SNR in the same region i , the region i which is represented by \mathcal{C}_i is further quantized with b_i bits. Therefore, the feedback bits of each user include two parts: one is the region index with B_C bits and the other is the quantization bits for that region.

IV. ANALYSIS OF SUM RATE LOSS

A natural question to ask is how many clusters (thresholds) should be set. Obviously, if many clusters are used, the number of region index bits B_C is increased. On the other hand, when a small number of clusters are used, the smallest threshold r_{c,N_c} becomes large. Then, the sum rate loss caused by the threshold r_{c,N_c} may increase. Therefore, an appropriate number of clusters is important for the design. We now

analyze the sum rate loss of the system caused by the smallest threshold r_{c,N_c} . We define a random variable Z_i as follows:

$$Z_i = \begin{cases} 0, & X_i > r_{c,N_c} \\ X_i, & X_i \leq r_{c,N_c} \end{cases}. \quad (16)$$

According to the maximum sum rate criterion, the exact SNR loss for the system is $Z_{(1)} = \max\{Z_1, Z_2, \dots, Z_K\}$. The probability of the rate loss event can be obtained by

$$P_L = P\{X_{(1)} \in (0, r_{c,N_c})\} = \prod_{i=1}^K (1 - \exp(-\lambda_i r_{c,N_c})). \quad (17)$$

Therefore, the sum rate loss can be expressed as

$$\Delta R = ME\{\log_2(1 + Z_{(1)})\}P_L. \quad (18)$$

Theorem 1. [15] *Let the means and variances of the random variables Z_1, Z_2, \dots, Z_K be $\mu = (\mu_1, \mu_2, \dots, \mu_K)$ and $\sigma^2 = (\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2)$, respectively. The closed form upper bounds on the expected value of the largest order statistic is:*

$$\begin{aligned} E\{Z_{(1)}\} &\leq \sum_{i=1}^K \left\{ \frac{\mu_i + \sqrt{(\mu_i - T)^2 + \sigma_i^2}}{2} \right\} + \frac{(2-K)T}{2} \\ &\triangleq \mu_{Z_{(1)}}^U \end{aligned} \quad (19)$$

where $T = \max_{1 \leq j \leq K} \{\mu_j + \frac{K-2}{2\sqrt{K-1}}\sigma_j\}$.

Applying Jensen's inequality and the upper bound of $E\{Z_{(1)}\}$, the sum rate loss on a certain beam can be bounded by

$$\begin{aligned} \frac{\Delta R}{M} &= E\{\log_2(1 + Z_{(1)})\}P_L \leq \log_2(1 + E\{Z_{(1)}\})P_L \\ &\leq \log_2(1 + \mu_{Z_{(1)}}^U)P_L. \end{aligned} \quad (20)$$

Using (20), for a given tolerable sum rate loss ΔR_U , the minimum number of clusters (thresholds) required can be determined.

Proposition 1. *When the total number of users approaches infinity, the sum rate loss caused by the smallest finite threshold r_{c,N_c} is negligible. Thus, the full multiuser diversity can be achieved.*

Proof: When the total number of users K approaches to infinity, $\lim_{K \rightarrow \infty} P_L = 0$. In addition, $Z_{(1)}$ is bounded by r_{c,N_c} . The sum rate loss on a certain beam becomes zero

$$\begin{aligned} \lim_{K \rightarrow \infty} \frac{\Delta R}{M} &= \lim_{K \rightarrow \infty} E\{\log_2(1 + Z_{(1)})\}P_L \\ &\leq \lim_{K \rightarrow \infty} \log_2(1 + E\{Z_{(1)}\})P_L \\ &\leq \lim_{K \rightarrow \infty} \log_2(1 + r_{c,N_c})P_L \\ &= 0. \end{aligned} \quad (21)$$

V. BIT ALLOCATION AND FEEDBACK LOAD ANALYSIS

A. Bit Allocation

Let $r_{c,0} = \infty$. Assume that the SNR region i $[r_{c,i}, r_{c,i-1})$, denoted $\mathfrak{C}_i, i = 1, 2, \dots, N_c$, is quantized with b_i bits. In region \mathfrak{C}_i , the quantization levels using b_i bits are expressed by $q_t^i, t = 1, 2, \dots, 2^{b_i}$, obtained by a pdf quantizer [16]. Thus, each level will occur with the same probability. Under the per user average feedback load constraint C_k for user k , the available bits will be assigned to the regions to maximize sum rate. Let $P_{\mathfrak{C}_i} = P\{X_{(1)} \in \mathfrak{C}_i\}, i = 1, 2, \dots, N_c$. The bit allocation problem can be described as follows:

$$\begin{aligned} \max_{(b_1, b_2, \dots, b_{N_c})} & M \sum_{i=1}^{N_c} \frac{P_{\mathfrak{C}_i}}{2^{b_i}} \sum_{t=1}^{2^{b_i}} \log(1 + q_t^i) \\ \text{s.t.} & \sum_{i=1}^{N_c} P_{\mathfrak{C}_i} b_i \leq C_k, k = 1, 2, \dots, K, \end{aligned} \quad (22)$$

where $P_{\mathfrak{C}_{k,j}} = P\{X_k \in \mathfrak{C}_j\} = e^{-\lambda_k r_{c,j}} - e^{-\lambda_k r_{c,j-1}}$. The closed-form integer solution for (22) does not exist. We use exhaustive search to find the optimal bit allocation set to maximize the sum rate.

B. Feedback Load Analysis

When a user's instantaneous SNR is smaller than the smallest threshold r_{c,N_c} , the user does not need to feedback. Thus, the feedback probability for user k is $e^{-\lambda_k r_{c,N_c}}$. The average total feedback load can be expressed as

$$F_b = M \sum_{k=1}^K \{e^{(-\lambda_k r_{c,N_c})} (B_C + C_k)\} \quad (23)$$

VI. SIMULATION RESULTS

In this section, we show the sum rate and feedback load performance for the proposed feedback scheme. The BS is equipped with $M = 4$ antennas and the total number of users is $K = 10 \sim 100$. In order to perform ZF beamforming, we let the number of receive antennas $N = 4$. The total transmit power P is 10W, while the additive white Gaussian noise power at the receivers σ_N^2 is 1W. Note that these numbers are selected only for illustration purpose. The elements of the channel matrix \mathbf{H}_i for the i th user are assumed to be i.i.d. complex Gaussian distribution with zero and variance σ_i^2 , where σ_i^2 are drawn uniformly from the interval $[0, 1]$ to model heterogeneous channels. The number of clusters in type-I and type-II schemes is set to four (thus $B_C = 2$) according to the tolerable sum rate loss $\Delta R_U = 10^{-2}$ bps/Hz. The feedback load constraint of user k is $C_k = 0.8, k = 1, 2, \dots, K$. In the simulation, the proposed feedback schemes are compared with the conventional scheme and the single-threshold schemes proposed in [9]. In the conventional feedback scheme, no matter what the instantaneous SNR is, it is always quantized with 3 bits. In the single-threshold scheme, the region (r_{th}, ∞) of the SNR is quantized using 3 bits. The single threshold r_{th} is established according to the scheduling outage probability P_{out} .

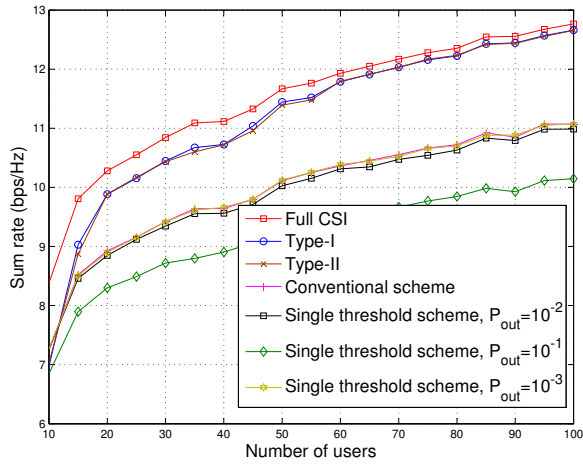


Fig. 2. Sum rate comparison between different feedback schemes.

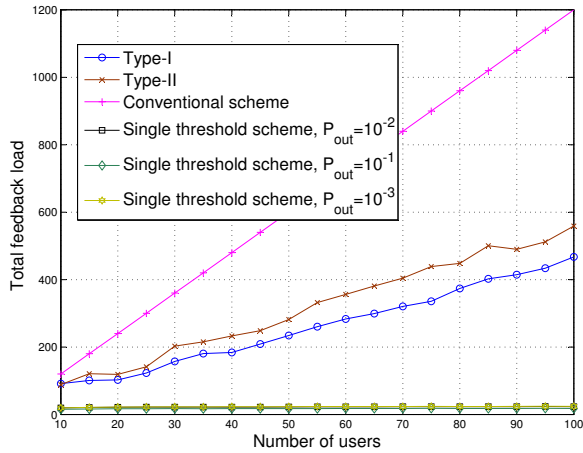


Fig. 3. Feedback load comparison between different feedback schemes.

In Fig. 2, the type-I feedback scheme achieves the highest sum rate, and the low-complexity type-II scheme achieves almost the same rate as the type-I scheme. As shown in Fig. 3, the feedback load of the conventional scheme increases linearly with the number of users. Using our proposed schemes, the total number of feedback bits can be dramatically reduced. Overall, the proposed schemes use fewer bits to achieve higher sum rate than the conventional scheme. The single-threshold scheme has lower feedback load, but suffers significantly in the sum rate performance.

In Fig. 4, we plot the sum rate vs. the total feedback load as an indication of the efficiency. The type-I scheme is the most efficient (i.e., making best use of the feedback bits), but has high computational complexity. The low-complexity type-II scheme not only achieves high sum rate but also reduces the feedback load significantly. The single-threshold feedback scheme only achieves a sum rate of about 11(bps/Hz) with a small feedback load.

VII. CONCLUSION

In this paper, we investigated the feedback load reduction problem in multiuser MIMO broadcast system. We proposed a

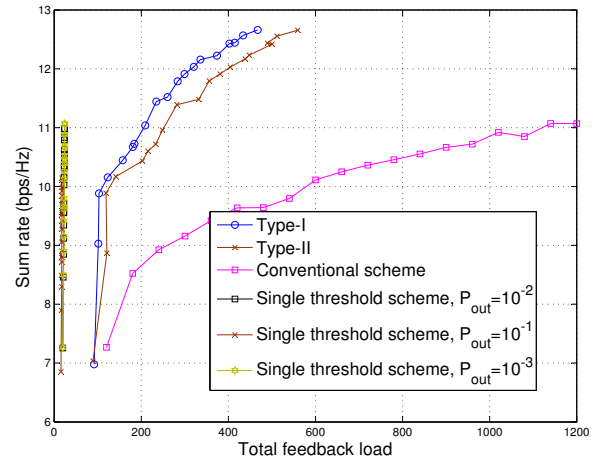


Fig. 4. The efficiencies of different feedback schemes.

cluster-based feedback scheme to reduce the feedback load in heterogenous and homogeneous channels. The bit allocation problem for the multiple-cluster feedback scheme was also discussed. The simulation results showed that, compared to the existing feedback schemes, the cluster-based feedback scheme can make the best use of the feedback bits to achieve good feedback load reduction while maintaining good sum rate performance.

REFERENCES

- [1] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. on Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.
- [2] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. on Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [3] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. on Inf. Theory*, vol. 50, pp. 1875–1892, 2004.
- [4] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 528–541, Mar. 2006.
- [5] N. Jindal, "MIMO Broadcast Channels With Finite-Rate Feedback," *IEEE J. Select. Areas Commun.*, vol. 52, no. 11, pp. 5045–5060, 2006.
- [6] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-Antenna Downlink Channels with Limited Feedback and User Selection," vol. 25, no. 7, pp. 1478–1491, Sep 2007.
- [7] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *IEEE Trans. on Inf. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.
- [8] M. Pugh and B. D. Rao, "Reduced Feedback Schemes Using Random Beamforming in MIMO Broadcast Channels," *IEEE Trans. on Signal Processing*, vol. 58, no. 3, pp. 1821–1832, March 2010.
- [9] D. Gesbert and M. S. Alouini, "How much feedback is multi-user diversity really worth?," *IEEE Int. Conf. on Commun.(ICC)*, vol. 1, pp. 234–238, June 2004.
- [10] J.-H. Li and H.-J. Su, "A Multi-threshold Scheme for Feedback Load Reduction in Multiuser MIMO Broadcast Channel," *IEEE PIMRC*, Sept. 2010.
- [11] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. on Inf. Theory*, vol. 45, pp. 139–157, 1999.
- [12] R. W. Heath, M. Airy, and A. J. Paulraj, "Multiuser diversity for MIMO wireless systems with linear receiver," in *proc. of the Asil. Conf. on Sig. Sys. and Comp.*, vol. 2, pp. 1194–1199, Nov. 2001.
- [13] D. A. Gore, R. W. Heath, and A. J. Paulraj, "Transmit Selection in Spatial Multiplexing Systems," *IEEE Communication Letter*, vol. 6, pp. 491–493, Nov. 2002.

- [14] T. Tang, R. W. Heath, S. Cho, and S. Yun, "Opportunistic Feedback for Multiuser MIMO Systems With Linear Receivers," *IEEE Trans. on Communications*, vol. 55, no. 5, May. 2007.
- [15] K. Natarajan D.Bertsimas and C.-P.Teo, "Tight bounds on expected order statistics," *Probab.Eng.Inform.Sc.*, vol. 20, no. 4, pp. 667–686, 2006.
- [16] J. L. Vicario, R. Bososio, and C. Anton-Haro, "A throughout analysis for opportunistic beamforming with quantized feedback," *IEEE PIMRC*, pp. 1–5, Sept. 2006.